

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-269224

(43)Date of publication of application : 09.10.1998

(51)Int.Cl.

G06F 17/28

G06F 17/27

G06F 17/30

(21)Application number : 09-076173

(71)Applicant : SUMITOMO METAL IND LTD

(22)Date of filing : 27.03.1997

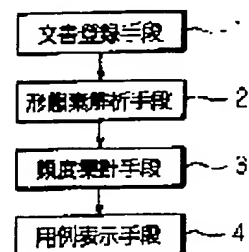
(72)Inventor : AOE JUNICHI
TSUDA KAZUHIKO
NAKAMURA MASAMI

(54) SYSTEM AND METHOD FOR ASSIGING ELECTRONIC DICTIONARY

(57)Abstract:

PROBLEM TO BE SOLVED: To make it possible to assist various operations for generating an electronic dictionary, by showing information needed to generate the electronic dictionary, i.e., information regarding the use frequencies and examples of respective words of a document.

SOLUTION: A document registering means 1 hierarchically classifies and registers documents, which are already published, according to proper parameters corresponding to their sources. A morpheme analyzing means 2 decomposes a registered document into words and converts the document into a sequence of morphemes consisting of word notation and parts of speech. A frequency totalizing means 3 totalizes the appearance frequencies of the morphemes of the registered document according to the analytic result of the morpheme analyzing means 2 by using the classification reference parameters of the document registering means 1 as units. An example display means 4 displays sentences before and after the morphemes so that how each specified morpheme is used in the actual document, can be found. Consequently, a proper example can be shown to a dictionary producer as the producer only desires.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-269224

(43) 公開日 平成10年(1998)10月9日

(51) Int.Cl. ⁸	識別記号	F I	
G 0 6 F 17/28		G 0 6 F 15/38	C
17/27			E
17/30		15/40	3 7 0 J
		15/401	3 1 0 D
		15/403	3 3 0 C
審査請求 未請求 請求項の数10 O L (全 7 頁) 最終頁に続く			

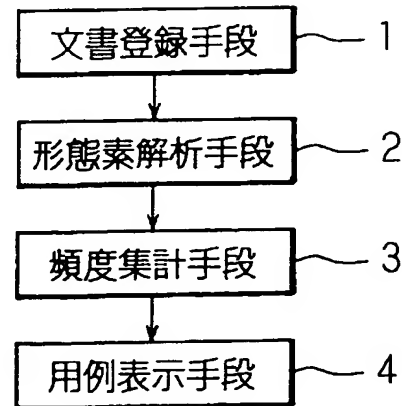
(21) 出願番号	特願平9-76173	(71) 出願人	000002118 住友金属工業株式会社 大阪府大阪市中央区北浜4丁目5番33号
(22) 出願日	平成9年(1997)3月27日	(72) 発明者	青江 順一 徳島県徳島市住吉2丁目7番36号
		(72) 発明者	津田 和彦 大阪府大阪市中央区北浜4丁目5番33号 住友金属工業株式会社内
		(72) 発明者	中村 雅巳 大阪府大阪市中央区北浜4丁目5番33号 住友金属工業株式会社内
		(74) 代理人	弁理士 河野 登夫

(54) 【発明の名称】 電子辞書作成支援システム及び電子辞書作成支援方法

(57) 【要約】

【課題】 電子辞書の作成作業に必要な情報を簡単に提示する電子辞書作成支援システムを提供する。

【解決手段】 既存の文書をその出典元を含む種々のパラメータに従って分類して登録する文書登録手段1と、登録された文書を単語表記と品詞とで構成する形態素の列に変換する形態素解析手段2と、各形態素の出現頻度を文書登録手段1における分類時のパラメ毎に集計する頻度集計手段3と、登録された文書におけるその用例を各形態素別に表示する用例表示手段4とを備え、電子辞書の作成作業の手助けとなる、文書における単語の出現頻度と指定した単語における用例との正確な情報を提供する。



【特許請求の範囲】

【請求項 1】 電子辞書の作成を支援するシステムであって、文書をその出典元を含む種々のパラメータに従って分類して登録する文書登録手段と、登録された文書を単語単位に分割し分割された単語の品詞を確定することにより、登録された文書を単語表記と品詞とで構成する形態素の列に変換する形態素解析手段と、各形態素の出現頻度を、前記文書登録手段における分類時のパラメータ毎に集計する頻度集計手段と、登録された文書におけるその用例を各形態素別に表示する用例表示手段とを備えることを特徴とする電子辞書作成支援システム。

【請求項 2】 登録する文書の中に含まれる同一内容異表記の複数の形態素または形態素列を 1 つの形態素または形態素列として集約する集約手段を更に備える請求項 1 記載の電子辞書作成支援システム。

【請求項 3】 前記形態素解析手段は、活用語の単語についてはその活用形も加えて形態素とし、前記頻度集計手段は、活用語の単語については表記が異なっても語幹が同一の単語は同一の形態素として集計すべくなした請求項 1 または 2 記載の電子辞書作成支援システム。

【請求項 4】 前記用例表示手段は、表示する用例の数を指定する手段を有する請求項 1 ～ 3 の何れかに記載の電子辞書作成支援システム。

【請求項 5】 前記頻度集計手段は、形態素の品詞別、活用語の形態素の活用形別、及び、前記集約手段で集約された形態素または形態素列別に、出現頻度を集計すべくなした請求項 3 または 4 記載の電子辞書作成支援システム。

【請求項 6】 電子辞書の作成を支援する方法において、文書をその出典元を含む種々のパラメータに従って分類して登録するステップと、登録された文書を単語単位に分割し分割された単語の品詞を確定することにより、登録された文書を単語表記と品詞とで構成する形態素の列に変換するステップと、各形態素の出現頻度を、前記文書登録手段における分類時のパラメータ毎に集計するステップと、登録された文書におけるその用例を各形態素別に表示するステップとを有することを特徴とする電子辞書作成支援方法。

【請求項 7】 登録する文書の中に含まれる同一内容異表記の複数の形態素または形態素列を 1 つの形態素または形態素列として集約するステップ更に有する請求項 6 記載の電子辞書作成支援方法。

【請求項 8】 前記形態素の列に変換するステップでは、活用語の単語についてはその活用形も加えて形態素とし、前記出現頻度を集計するステップでは、活用語の単語については表記が異なっても語幹が同一の単語は同一の形態素として集計する請求項 6 または 7 記載の電子辞書作成支援方法。

【請求項 9】 前記用例を表示するステップでは、表示する用例の数を指定する請求項 6 ～ 8 の何れかに記載の

電子辞書作成支援方法。

【請求項 10】 前記出現頻度を集計するステップでは、形態素の品詞別、活用語の形態素の活用形別、及び、集約された形態素または形態素列別に、出現頻度を集計する請求項 8 または 9 記載の電子辞書作成支援方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、言語をコンピュータ解析するために必要な電子辞書の作成を支援するシステムに関し、特に、電子辞書の編集に必要な情報を提供する電子辞書作成支援システム及び方法に関する。

【0002】

【従来の技術】ある言語をコンピュータによって解析するための形態素解析プログラム、構文解析プログラム、意味解析プログラム等で用いる電子辞書を作成する場合、従来では、その言語の専門家が、言語解析プログラムの解析結果を見て、自身の知識と経験とに基づいて、手作業で電子辞書の編集を行っている。

【0003】

【発明が解決しようとする課題】このようなシステムを用いない従来の手法では、以下に述べるような問題点がある。登録単語の選定作業においては、辞書編集者自身の好みまたは偏見が入り、使用頻度が高い語を正確に選定することが困難である。選定した登録単語の構文構造及び意味内容の登録作業においては、参考とする数多くの文書を使用法の偏りがないように揃えることは困難であり、多くの時間が浪費されると共に、参考例の不足による構文構造または意味内容の記述不備が生じる可能性が高い。また、従来では、専門知識を熟知した人のみが、その個人的な判断で電子辞書を編集するので、生産性が悪いだけでなく、専門的な知識身につけるのに多大な時間が必要となるため、多くの人材を育成することは不可能である。

【0004】本発明は斯かる事情に鑑みてなされたものであり、電子辞書の作成に必要な情報、つまり文書における各単語の使用頻度及び各単語の用例についての情報を提示することにより、電子辞書の作成時の様々な作業の手助けとなる電子辞書作成支援システム及び方法を提供することを目的とする。

【0005】

【課題を解決するための手段】請求項 1 に係る電子辞書作成支援システムは、電子辞書の作成を支援するシステムであって、文書をその出典元を含む種々のパラメータに従って分類して登録する文書登録手段と、登録された文書を単語単位に分割し分割された単語の品詞を確定することにより、登録された文書を単語表記と品詞とで構成する形態素の列に変換する形態素解析手段と、各形態素の出現頻度を、前記文書登録手段における分類時のパラメータ毎に集計する頻度集計手段と、登録された文

書におけるその用例を各形態素別に表示する用例表示手段とを備えることを特徴とする。

【0006】請求項2に係る電子辞書作成支援システムは、請求項1において、登録する文書の中に含まれる同一内容異表記の複数の形態素または形態素列を1つの形態素または形態素列として集約する集約手段を更に備えることを特徴とする。

【0007】請求項3に係る電子辞書作成支援システムは、請求項1または2において、前記形態素解析手段は、活用語の単語についてはその活用形も加えて形態素とし、前記頻度集計手段は、活用語の単語については表記が異なっても語幹が同一の単語は同一の形態素として集計すべくなしたことを特徴とする。

【0008】請求項4に係る電子辞書作成支援システムは、請求項1～3の何れかにおいて、前記用例表示手段は、表示する用例の数を指定する手段を有することを特徴とする。

【0009】請求項5に係る電子辞書作成支援システムは、請求項3または4において、前記頻度集計手段は、形態素の品詞別、活用語の形態素の活用形別、及び、前記集約手段で集約された形態素または形態素列別に、出現頻度を集計すべくなしたことを特徴とする。

【0010】請求項6に係る電子辞書作成支援方法は、電子辞書の作成を支援する方法において、文書をその出典元を含む種々のパラメータに従って分類して登録するステップと、登録された文書を単語単位に分割し分割された単語の品詞を確定することにより、登録された文書を単語表記と品詞とで構成する形態素の列に変換するステップと、各形態素の出現頻度を、前記文書登録手段における分類時のパラメータ毎に集計するステップと、登録された文書におけるその用例を各形態素別に表示するステップとを有することを特徴とする。

【0011】請求項7に係る電子辞書作成支援方法は、請求項6において、登録する文書の中に含まれる同一内容異表記の複数の形態素または形態素列を1つの形態素または形態素列として集約するステップ更に有することを特徴とする。

【0012】請求項8に係る電子辞書作成支援方法は、請求項6または7において、前記形態素の列に変換するステップでは、活用語の単語についてはその活用形も加えて形態素とし、前記出現頻度を集計するステップでは、活用語の単語については表記が異なっても語幹が同一の単語は同一の形態素として集計することを特徴とする。

【0013】請求項9に係る電子辞書作成支援方法は、請求項6～8の何れかにおいて、前記用例を表示するステップでは、表示する用例の数を指定することを特徴とする。

【0014】請求項10に係る電子辞書作成支援方法は、請求項8または9において、前記出現頻度を集計するス

テップでは、形態素の品詞別、活用語の形態素の活用形別、及び、集約された形態素または形態素列別に、出現頻度を集計することを特徴とする。

【0015】図1は、本発明の電子辞書作成支援システムの構成を示す図であり、本発明のシステムは、図1に示すように、文書登録手段1と、形態素解析手段2と、頻度集計手段3と、用例表示手段4とを有する。

【0016】文書登録手段1は、新聞、小説、随筆、解説文、論文等で既に出版されている文書を、例えば新聞ではその発行年月、発行社、掲載欄等、小説ではそのジャンル、作者、年代別等、その出典元に応じた適当なパラメータにより階層的に分類して登録する。このようにしておけば、後述する形態素の集計結果の獲得及び用例の表示を、分類したパラメータ毎に行える。

【0017】形態素解析手段2は、登録された文書を単語に分解し、その単語表記と品詞とで構成される複数の形態素の列に文書を変換する。この際、活用がある単語、つまり活用語（具体的には用言及び助動詞）については、文書中における活用形も確定しておけば、後述する形態素の集計結果及び用例表示において、品詞だけでなく活用形も考慮することができる。

【0018】頻度集計手段3は、形態素解析手段2での解析結果に基づいて、文書登録手段1における分類基準であるパラメータを単位として、登録された文書における形態素の出現頻度を集計する。このようにして、各パラメータ毎の形態素の出現頻度のデータが簡単に得られる。この際、活用語の形態素については、表記が異なっても、つまり活用形が異なっても、語幹が同じである単語を含む形態素は同一の形態素として集計するようにすれば、より現実的な形態素の集計結果が得られる。また、形態素の品詞別及び活用語の形態素の活用形別に、出現頻度を集計することも可能である。

【0019】用例表示手段4は、指定された形態素別に、その形態素が実際の文書でどのように使われているかが分かるように、その形態素の前後文を表示する。この際、文書登録手段1における分類基準であるパラメータを単位として、1つの形態素毎に全ての用例を表示することも可能であると共に、表示する用例の数を指定することも可能である。このようにすることにより、辞書制作者に対して、その人が所望するだけの適切な用例を提示することができる。

【0020】なお、以上のような構成に加えて、表記は異なるが同一の意味内容を持つ形態素または形態素列については、同一の形態素または形態素列として集約する集約手段を備えておけば、例えば、異表記同義語の単語を同一の単語とした場合の形態素の集計結果、異表記同義語の単語は同一の単語としない場合の形態素の集計結果など、様々な条件での形態素の集計結果を得ることができる。

【0021】本発明の電子辞書作成支援システムでは、

以上のように、既存の文書における各単語の使用頻度、各単語の使用例等、電子辞書の編集に必要な情報を、過不足なく適切に辞書作成者に提供することができる。よって、電子辞書の作成作業の支援に大いに役立つことになる。

【0022】

【発明の実施の形態】本発明を、その実施の形態を示す図面を参照して具体的に説明する。なお、以下の説明では、日本語電子辞書の作成を支援するシステムを例として詳述する。

【0023】図2は、本発明の日本語電子辞書作成支援システムを実現するハードウェアの構成を示す図であり、コンピュータ11と、形態素の出現頻度の集計結果、形態素の用例などを表示するディスプレイ12と、コンピュータ11に接続されたキー入力を行うキーボード13及び指示入力を行うマウス14と、磁気ディスク、光ディスク等の記録媒体15と、登録した日本語文書、形態素の分析結果などを記憶するデータベースとしての外部記憶手段16とを備える。本発明は、後述する各処理を含むコンピュータプログラムを記録してある記録媒体15をコンピュータ11のディスクドライブ11aに装着し、記録媒体15に記録されたコンピュータプログラムをロードすることにより後述する各処理を実行する。

【0024】図3は、本発明の日本語電子辞書作成支援システムの構成を示す図であり、本発明の日本語電子辞書作成支援システムは、前述した図1と同様の、既存の日本語文書を種々のパラメータに従って分類して登録する文書登録手段1、登録された文書を複数の形態素の列に変換する形態素解析手段2、登録された文書における形態素の出現頻度を集計する頻度集計手段3、及び、形態素毎にその形態素の用例を表示する用例表示手段4に加えて、同一内容の複数の形態素または形態素列を1つの形態素または形態素列に集約する集約手段5とを備える。

【0025】次に、動作について説明する。図4はその手順を示すフローチャートである。まず、文書登録手段1により、新聞、小説、随筆、解説文、論文等既に出版されている処理対象の既存の日本語文書を分類して登録する(S1)。この際、例えば新聞ではその発行年月、発行社、掲載欄等のパラメータによって紙面の文書を分類し、小説ではそのジャンル、筆者、年代別等のパラメータにより文書を分類する。このようにしておくことにより、簡単な操作によって、形態素の集計結果及び用例文の表示を、分類したパラメータ毎に見ることができると共に、形態素の出現頻度、個別形態素の利用方法及び形態素の意味などにおける各パラメータ間での差を比較することもできる。

【0026】そして、形態素解析手段2により、登録された文書を単語に分解して、形態素(単語表記+品詞+活用形(活用語のみ))の列を形成する(S2)。例え

ば「赤い花が咲く」という文書を例にした場合、「赤い」「花」「が」「咲く」という各単語に分解されて、図5に示すような形態素の列になる。このような処理は、登録されているすべての文書に対して実行する。

【0027】次に、頻度集計手段3により、形態素解析手段2での解析結果に基づいて、文書登録手段1における分類基準であるパラメータを単位として、形態素の出現頻度を集計する(S3)。このように、文書登録手段1での分類基準のパラメータを単位とする頻度の集計結果が得られるので、例えば新聞と小説との間における、形態素の出現頻度の差、同じ形態素の使用方法及び意味の差などを容易に比較することができる。

【0028】また、頻度集計手段3では品詞別に集計結果を得る。例えば、「愛」という表記で集計する場合、「愛は永遠だ」の「愛」は普通名詞であり、「愛すること」の「愛」はサ変動詞の語幹となる。このように「愛」で集計すると2種類の品詞を持つ「愛」が集計されるので、その出現頻度を求める際には、普通名詞の「愛」の頻度、サ変動詞の「愛」の頻度というように個別に集計する。

【0029】更に、活用語、具体的には用言(動詞、形容詞、形容動詞)または助動詞である形態素については、文書中に出ているのが如何なる活用形であっても1つの形態素として取り扱って、その出現頻度を集計する。例えば、「咲く」というカ行5段活用動詞を考えた場合、その活用形は「咲か」または「咲こ」(未然形)、「咲き」または「咲い」(連用形)、「咲く」(終止形)、「咲く」(連体形)、「咲け」(假定形)、「咲け」(命令形)となるが、これらはどれも「咲」という語幹が同じであって活用形が異なるだけであるので、同一の形態素と見做して出現頻度を集計する。なお、その集計結果の内訳として各活用形における出現頻度の集計結果も求めておく。

【0030】集約手段5により、表記は異なるが同一の意味内容を持つ複数のものを1つに集約する(S4)。集約手段5では、集約する語の指定方法として文字単位、形態素単位及び形態素列単位の3種類の指定方法が可能である。集約手段5は、図6に示すように、内部にこれらの集約を実施する際の集約規則を格納しており、この集約規則に従って、表記は異なるが内容は同一である複数の語が1つの語に集約される。得られた集約結果は、頻度集計手段3へ送られる。

【0031】まず、文字単位の集約規則に従って、文字単位での集約が行える。文字単位の指定では、例えば、「沢」と「澤」とを同一の文字として指定すれば、「沢田」と「澤田」という指定した文字のみが異なる2つの形態素が同一の形態素として集約できる。従って、新旧の文字、当て字等、表記は異なるが意味は同じである文字を同一の文字として集約でき、また、例えばJISコード等、コンピュータコードに登録されていない文字も

コンピュータコードに登録されている文字と同一の文字として取り扱うことができ、より現実に即した対応が可能である。なお、文字単位指定において同一文字として取り扱う文字の数は3つ以上も可能であり、その数に制限はない。

【0032】形態素単位での集約には、表記の集約規則に従う表記による集約機能と、品詞の集約規則に従う品詞による集約機能と、属性（意味、対応英単語など）の集約規則に従う属性による集約機能が存在する。表記による集約では、表記は異なるが同一の意味内容を持つ単語、即ち異表記同義語については、同一の単語として集約する。例えば「行う」と「行なう」とでは送りがなの違いだけであり、「コンピュータ」と「コンピューター」とでは語尾を伸ばすか伸ばさないかの違いだけであるので、このようなものは1つの単語として集約する。品詞による集約では、同一の品詞である語を集約する。例えば、「感嘆詞」という品詞を指定した場合に、「ああ」、「えっ」、「さあ」などのすべての感嘆詞の語を集約する。属性による集約では、各形態素の属性として英単語を登録していた場合、例えば「paper」という属性を持つ形態素には「紙」、「新聞」、「論文」などがあるので、「paper」という英単語属性を指定することにより、これらの「紙」、「新聞」、「論文」などの語を集約する。

【0033】形態素列単位の指定では、個々の形態素の指定方法は形態素単位での指定と同様に表記、品詞及び属性で指定することも可能であると共に、各形態素毎の指定方法の混在も可能であり、また、同等として取り扱う形態素列の形態素の数の相違も可能としている。このような、形態素列単位の指定による集約は、形態素列単位の集約規則に従う。例えば、「受験するための資格」、「入場するための許可」は、それぞれ「受験資格」、「入場許可」という複合語を用いることによって短縮表現が可能である形態素列である。これらの形態素列を集約する際には、「サ変（品詞）」＋「するための（表記）」＋「名詞（品詞）」＝「サ変（品詞）」＋「名詞（品詞）」というように集約できる。

【0034】図7は、集約手段5におけるこのような形態素列の指定をダイアログ形式で行える一例を示す図である。図において、51は指定した形態素列を表示する表示枠であり、1形態素列が1行で表示される。52は各形態素の指定方法を選択する選択ボックスであって「表記」、「品詞」、「属性」の何れかを選択する。また、ボタン53は形態素毎の指定完了を決定する際に押されるボタンであり、ボタン54は形態素列の完了を決定する際に押されるボタンである。

【0035】まず、ユーザは選択ボックス52で形態素列の要素となる形態素を指定する。選択ボックス52の左側のオン・オフスイッチ52aにより「表記」、「品詞」、「属性」の何れで指定するかを決定する。指定する項目

を決定した後、更に指定した項目をスイッチ52aの右の四角枠52bに入力する。表記のための枠は、テキスト入力機能となっていて、任意の文字列をその枠52bに入力する。品詞のための枠52cは、プルダウンメニューで指定可能な品詞が表示されるようになっており、その中から品詞を指定する。属性のための枠52dも、プルダウンメニューで指定可能な属性が表示されるようになっており、所望の属性を選択した後、その下の枠52e内に任意の指定要素を入力する。例えば、属性として「英単語属性」を枠52d内で選択し、下の枠52eには「paper」と入力する。以上のようにして形態素の指定が完了すると、ボタン53を押すことによって指定した形態素が表示枠51に表示される。このような操作を繰り返し、最後にボタン54を押すことによって形態素列が指定表示される。

【0036】頻度集計手段3には、集約手段5から上述したような種々の集約結果が送られる。そして、これらの集約結果を形態素の出現頻度に反映させたい場合には（S5：YES）、それらの集約結果を参考にして、様々な種類の出現頻度の集計結果を得ることができる（S6）。

【0037】用例表示手段4により、指定された形態素または形態素列の前後の文を表示する（S7）。用例表示手段4では、登録されたすべての文書における用例を表示することも可能であり、また、文書登録手段1での分類時のパラメータを単位としてすべての文書における用例を表示することも可能であり、更に、集約手段5で集約された表記は異なるが同一の意味を持つ形態素及び形態素列の用例を表示することも可能である。

【0038】出現頻度が高い形態素についてはそのすべての用例を見ることは困難であるので、その表示数を絞り込む機能を用例表示手段4は有している。この場合には、文書登録手段1で分類したパラメータ別の出現頻度、及び、集約手段5での集約結果を考慮した形で絞り込みを行うこととし、絞り込みによって偏った用例ばかりが表示されることが無いようにする。例えば、「沢田」と「澤田」とにおける出現頻度が図8（a）で示されているような場合、合計200個が出現しているのですべての用例を一目で見るとは困難である。よって、表示する用例の数を40個（全体の1／5）と指定することにより、表示するそれぞれの用例の数もこれに合わせて、図8（b）に示すように1／5ずつに絞り込む。このようにすることにより、各パラメータでの出現頻度、及び、集約手段5での集約結果を反映させた状態で指定された数だけの用例を表示することができ、日本語電子辞書の作成者が所望する適切で適正量の用例を提示することが可能となる。

【0039】なお、上述した例では、日本語電子辞書の作成を支援するシステムの場合について説明したが、日本語以外の他の言語の電子辞書の作成を支援する場合に

も、本発明を同様に適用できることは勿論である。

【0040】

【発明の効果】以上のように、本発明では、電子辞書の編集作業に必要な情報、つまり、文書における単語の出現頻度と指定した単語における用例との正確な情報を、集約機能により整理された形式で、電子辞書の作成者に提供することが可能になる。この結果、従来では不可能であったような大量の文書を参考にすることが可能となり、より正確で現状の言語状況に合った電子辞書を作成できる。このように、本発明は、電子辞書の作成支援に

【図面の簡単な説明】

【図1】本発明の電子辞書作成支援システムの構成図である。

【図2】本発明の日本語電子辞書作成支援システムを実現するハードウェアの構成図である。

【図3】本発明の日本語電子辞書作成支援システムの実

施の形態の構成図である。

【図4】本発明の日本語電子辞書作成支援方法の動作手順を示すフローチャートである。

【図5】本発明における形態素列の一例を示す図である。

【図6】本発明における集約手段の集約規則の格納例を示す図である。

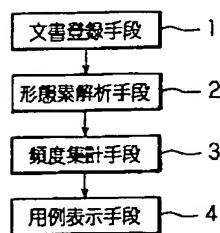
【図7】本発明における集約手段での形態素列の指定をダイアログ形式で行える一例を示す図である。

【図8】本発明における用例表示の絞り込みの例を示す図である。

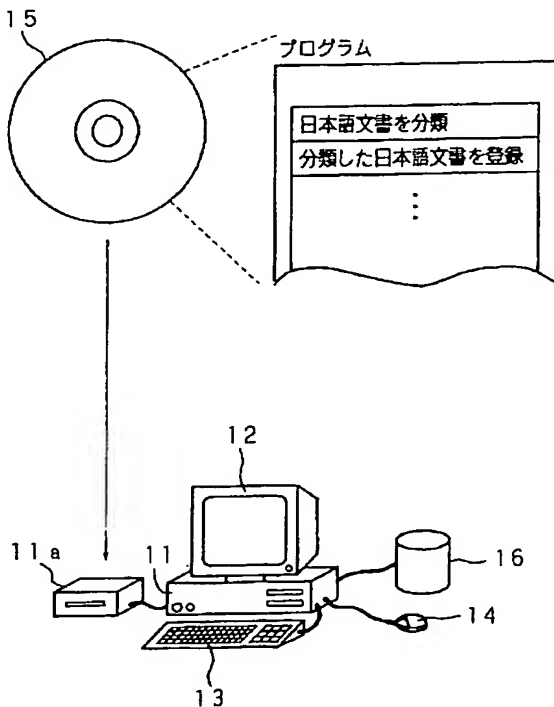
【符号の説明】

- 1 文書登録手段
- 2 形態素解析手段
- 3 頻度集計手段
- 4 用例表示手段
- 5 集約手段

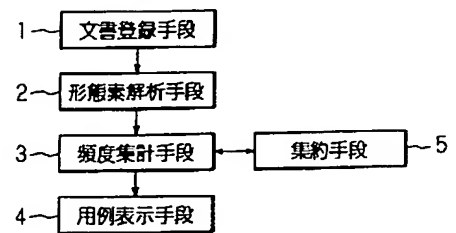
【図1】



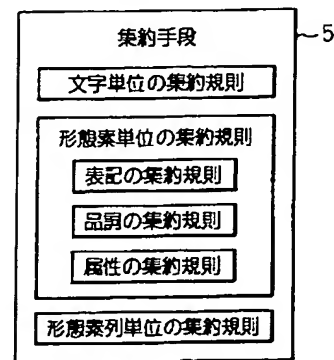
【図2】



【図3】



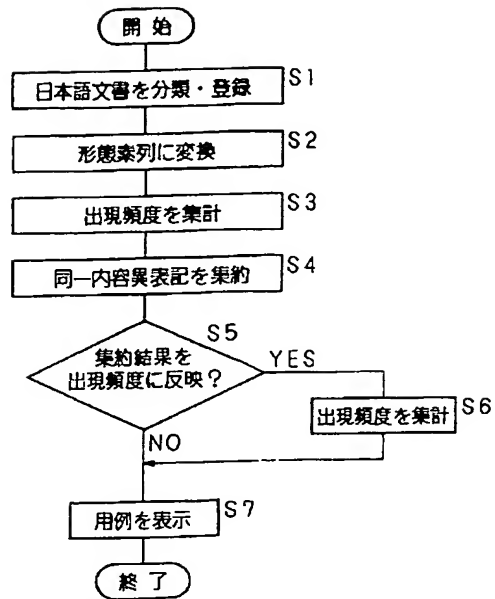
【図6】



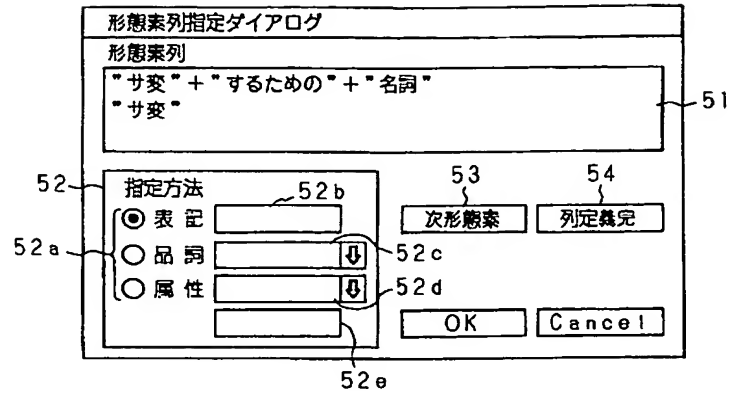
【図5】

赤い	花	が	咲く
形容詞	普通名詞	格助詞	5段動詞
連体形			終止形

【図4】



【図7】



【図8】

(a)

	沢田	澤田
パラメータA	50	20
パラメータB	100	30

(b)

	沢田	澤田
パラメータA	10	4
パラメータB	20	6

フロントページの続き

(51) Int. Cl.⁶

識別記号

F I

G O 6 F 15/403

3 4 0 B